Proceedings of the Twenty-ninth (2019) International Ocean and Polar Engineering Conference
Honolulu, Hawaii, USA, June 16-21, 2019

ISBN 978-1 880653 85-2; ISSN 1098-6189

www.isope.org

Intercomparison of Wave Reanalysis based on ERA5 and WW3 Databases

Christos Stefanakos
Dept. of Environment and New Resources, SINTEF Ocean
Trondheim, Norway

ABSTRACT

In this work an intercomparison of two wave reanalysis products, namely ERA5 and CFSR-W (WW3 hereafter), is presented. The datasets are assessed by means of several statistical features, such as seasonal variability, quantiles of the probability distribution, annual and inter-annual variability, and several error metrics. Both datasets cover a period of 31 years (1979-2009), a fact that assures that most of the long-scale features are equally present in both datasets. The analysis performed is depicted both on a global and regional basis. The results are also assessed on the basis of a global satellite altimeter dataset.

KEY WORDS: significant wave height; wave climatologies; ERA5; WW3; seasonal analysis; probability analysis

INTRODUCTION

The study of ocean wave climate is of great importance for a large number of applications, including among others: ocean climate change studies, design of ships and offshore/coastal structures, planning of sea operations, wind and wave energy conversion.

There various sources of wave data, namely (i) visual observations (Gulev and Grigorieva, 2006), (ii) satellite altimeters (Lefèvre and Cotton, 2001; Young et al., 2011), (iii) *in situ* buoy measurements, and (iv) numerical wave models (The WAMDI Group, 1988; Tolman et al., 2002). Each one of them have its own advantages and disadvantages. The great advantage of numerical models is their wide coverage on a high time and space resolution in a global scale, making it possible to produce long-term wind and wave climatologies.

The quality of wave data is heavily dependent on the quality of the wind forcing, improvements on the physical modelling and assimilation techniques among others. At least two meteorological centers (NCEP and ECMWF) work constantly the last decades towards the improvement of their forecast models. For this, apart from the results of their operational versions, they perform from time to time reanalysis studies to offer homogenized wave products.

Stopa (2018, Table 1) summarizes most of the well known wind datasets used throughout the years to generate global wave hindcasts. These wind products led to several global reanalyses of wave datasets the last decades. Among them, it is worth noted ERA-15, ERA-40, ERA-Interim from ECMWF (Sterl and Caires, 2005; Rascle and Ardhuin, 2013), NCEP/NCAR, CSFR-W from NCEP (Cox and Swail, 2001; Chawla et al., 2013; Rascle and Ardhuin, 2013). Other climatologies include HIPOCAS (Guedes Soares et al., 2002), GOW1 and 2 (Reguero et al., 2012; Perez et al., 2017).

The emergence of these reanalysis databases has been followed by intercomparison studies against each other and/or against other sources of data such as buoy and satellite measurements. For example, Caires et al. (2004) have compared six different reanalysis datasets (ERA-40 and NCEP/NCAR among them) against NOAA buoy and TOPEX/Poseidon altimeter datasets. Semedo et al. (2011) compared ERA-40 against visual observations and satellite data. Stopa in a series of works has analyzed CSFR-W wave data and their relation to climate indices (Stopa et al., 2013), against ERA-Interim and a number of altimeter data (Stopa and Cheung, 2014), using 12 different wind fields as input and against two different satellite datasets (Stopa, 2018). Finally, (Campos and Guedes Soares, 2016) have compared the results of HIPOCAS climatology against CSFR-W and ERA-Interim.

In the present work, an intercomparison of two wave databases is performed: a) the CFSR-W (Chawla et al., 2013; Rascle and Ardhuin, 2013) called for simplicity WW3, and b) the newly released ERA5 by ECMWF through Copernicus Climate Data Store (Copernicus Climate Change Service (C3S), 2017). The former has been extensively studied by Stopa and coauthors (Stopa et al., 2013; Stopa and Cheung, 2014), whereas the latter is studied here for the first time. In the next section, a brief description of the two datasets is given, as well as a description of the analysis procedure followed. Then, numerical results are presented and commented and finally conclusions are drawn. It should be noted that, although wind speed statistics are also available, they have been omitted from the present work due to space limitation and can be provided upon request.

METHODOLOGY

Data Used

In the present study, three datasets have been used for the intercomparison, covering the entire globe. The two of them have been generated by means of third-generation numerical wave models (WAVEWATCH III and WAM), and the third one consist of satellite altimeter measurements merged from several satellite missions. The model data are in regular gridded netcdf format (361 lats×720 lons = 259920 datapoints), while the satellite data are stored in tracks.

The first dataset, called hereafter for simplicity WW3, consists of fields of significant wave height covering the entire globe for the period 1979-2009 in 3-h intervals (31 years × 2920 3-h = 90520 time instances). For a more detailed description of the data, one can see here: `https://polar.ncep.noaa.gov/waves/hindcasts/nopp-phase2.php`.

The second dataset, ERA5, which also consists of fields of significant wave height, has recently been released. Although the data cover the period 1979-2018 in hourly intervals, only the period 1979-2009 in 3-hourly intervals has been taken into account. For a more detailed description, see: `https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5`.

In addition, and for comparison purposes, satellite altimeter data from the archive of IFREMER have been used. The archive contains data from nine altimeter missions, namely ERS-1,2, ENVISAT, TOPEX/Poseidon, Jason-1,2, GEOSAT-FO, Cryosat-2, SARAL, covering the period 1992-2016. More detailed information about the missions, as well as about the validation against buoy data and the induced corrections can be found in Queffeulou and Croizé-Fillon (2017).

Since satellite data have a different time-space data structure, and in order to make possible a comparison between them and the model data, mean monthly values are calculated for both sources for 13 different subregions; see Fig. 1. These discrete non-overlapping subregions have been defined by Alves (2006), such that the wave conditions within each of them to be similar.

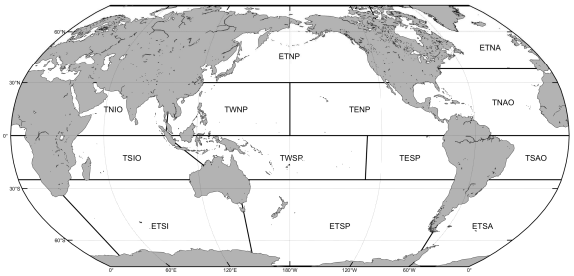

Fig. 1. Selected subregions of the oceans: Extratropical South Indian (ETSI), Extratropical South Pacific (ETSP), Extratropical South Atlantic (ETSA), Tropical South Indian Ocean (TSIO), Tropical Western South Pacific (TWSP), Tropical Eastern South Pacific (TESP), Tropical South Atlantic Ocean (TSAO), Tropical North Indian Ocean (TNIO), Tropical Western North Pacific (TWNP), Tropical Eastern North Pacific (TENP), Tropical North Atlantic Ocean (TNAO), Extratropical North Pacific (ETNP), Extratropical North Atlantic (ETNA)

Statistical Analysis Procedure

The data are available as fields in the form

$$\{X(t_i, \phi_j, \lambda_k), \quad i = 1, \ldots, I, \; j = 1, \ldots, J, \; k = 1, \ldots, K\}, \tag{1}$$

where i runs over the time instances, and j, k over the latitudes and longitudes, respectively.

In the sequel, two kind of analysis are performed: a) *field analysis*, showing results for the entire field, and b) *synoptic analysis*, showing results averaged over all (or a subset of datapoints) datapoints. In this way, various statistics related to different aspects of the datasets are depicted.

First, the time index is reparametrized according to Buys-Ballot triple index (Stefanakos et al., 2006) in order to properly treat variability at different time scales. Hence, the following triple index (y, m, n) will be used. The first component y is the yearly index. The second m=$\{1, 2, \ldots, 12\}$ is the monthly index. The third n=$\{1, 2, \ldots, N_m\}$ represents the time within a month, whith N_m be the number of 3-hourly observations within the m-th month.

According to the three-index notation, time series $X(t_i)$ at a location (ϕ_j, λ_k) is reindexed as follows:

$$\{X(y, m, n, \cdot, \cdot), \;\; y = 1, \ldots, \;\; m = 1, \ldots, 12, \;\; n = 1, \ldots, N_m\}. \tag{2}$$

The three indices (y, m, n) represent three different time scales, making it possible to explicitly define statistics with respect to each one of them, separately.

Seasonal Analysis. First, the fields of monthly values of mean value and standard deviation are formed

$$\mu_3(y, m, \cdot, \cdot) = \frac{1}{N_m} \sum_{n=1}^{N_m} X(y, m, n, \cdot, \cdot), \tag{3}$$

$$\sigma_3(y, m, \cdot, \cdot) = \sqrt{\frac{1}{N_m} \sum_{n=1}^{N_m} \left[X(y, m, n, \cdot, \cdot) - \mu_3(y, m, \cdot, \cdot)\right]^2}. \tag{4}$$

Then, the mean monthly values are obtained by averaging Eqs. 3~4 over the years Y:

$$\widetilde{\mu}_3(m, \cdot, \cdot) = \frac{1}{Y} \sum_{y=1}^{Y} \mu_3(y, m, \cdot, \cdot), \tag{5}$$

$$\widetilde{\sigma}_3(m, \cdot, \cdot) = \frac{1}{Y} \sum_{y=1}^{Y} \sigma_3(y, m, \cdot, \cdot), \quad m = 1, 2, \ldots, 12. \tag{6}$$

These parameters are also known as seasonal mean value and seasonal standard deviation, depicting the seasonal patterns of the data, and they have been used in a nonstationary time series modelling suitable for metocean and maritime parameters; see Athanassoulis and Stefanakos (1995); Stefanakos et al. (2006); Stefanakos and Schinas (2014).

If, further, one average over all ϕ_j and λ_k, a synoptic picture for the mean monthly values in Eq. 5~6 is obtained

$$\widetilde{\mu}_3(m) = \frac{1}{J\,K} \sum_j \sum_k \widetilde{\mu}_3(y, m, \cdot, \cdot), \tag{7}$$

$$\widetilde{\sigma}_3(m) = \frac{1}{J\,K} \sum_j \sum_k \widetilde{\sigma}_3(m, \cdot, \cdot). \tag{8}$$

Combining Eqs. 5~6, or equiv. Eqs. 7~8, one can calculate the coefficient of variation

$$\widetilde{cv}_3(m, \cdot, \cdot) = \frac{\widetilde{\sigma}_3(m, \cdot, \cdot)}{\widetilde{\mu}_3(m, \cdot, \cdot)}, \tag{9}$$

depicting the Mean Monthly Variability (MMV) of the field (equiv. of the averaged data).

Similarly, if, instead of Eqs. 3~4, one calculates the yearly values of

mean value and standard deviation

$$\mu_{32}(y,\cdot,\cdot) = \frac{1}{M\,N_m}\sum_{m=1}^{M}\sum_{n=1}^{N_m} X(y,m,n,\cdot,\cdot), \tag{10}$$

$$\sigma_{32}(y,\cdot,\cdot) = \sqrt{\frac{1}{M\,N_m}\sum_{m=1}^{M}\sum_{n=1}^{N_m}[X(y,m,n,\cdot,\cdot)-\mu_{32}(y,\cdot,\cdot)]^2}, \tag{11}$$

and, then, applies again Eq. 9, the Mean Annual Variability (MAV) is obtained as

$$\widetilde{cv}_{32}(\cdot,\cdot) = \frac{\widetilde{\sigma}_{32}(\cdot,\cdot)}{\widetilde{\mu}_{32}(\cdot,\cdot)}, \tag{12}$$

where $\widetilde{\mu}_{32}(\cdot,\cdot)$ and $\widetilde{\sigma}_{32}(\cdot,\cdot)$ are obtained from Eqs. 5~6.

Further, the year-to-year variability, or Inter-Annual Variability (IAV), is also defined as (Stopa et al., 2013)

$$\widehat{cv}_{32}(\cdot,\cdot) = \frac{\widehat{\sigma}_{32}(\cdot,\cdot)}{\widehat{\mu}_{32}(\cdot,\cdot)} \tag{13}$$

with

$$\widehat{\mu}_{32}(\cdot,\cdot) = \frac{1}{Y}\sum_{y=1}^{Y}\mu_{32}(y,\cdot,\cdot), \tag{14}$$

$$\widehat{\sigma}_{32}(\cdot,\cdot) = \sqrt{\frac{1}{Y}\sum_{y=1}^{Y}[\mu_{32}(y,\cdot,\cdot)-\widehat{\mu}_{32}(\cdot,\cdot)]^2}. \tag{15}$$

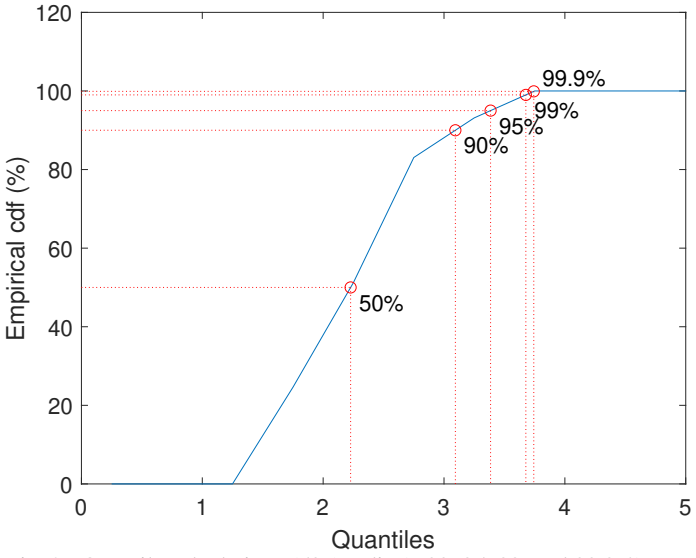

Fig. 2. Quantile calculations (50 (median), 90, 95, 99, and 99.9%)

Probability Analysis. If the probability function of month m and year y is denoted by

$$\mathcal{F}(x;y,m,\cdot,\cdot), \quad y = 1,2,\ldots,Y, \quad m = 1,2,\ldots,M, \tag{16}$$

then, the quantiles of it are given by

$$x_p = \mathcal{F}^{-1}(p;y,m,\cdot,\cdot), \quad p \in [0,1], \tag{17}$$

In the present paper the quantiles of $p = 50\%$ (median), 90%, 95%, 99%, and 99.9% are calculated; see also Fig. 2.

The overall probability function, and consequently the associated quantiles, can be obtained by adding the frequencies of each month m and year y. Similarly, in the synoptic analysis, the probability is obtained by adding the frequencies of all points together.

Error Analysis. The following statistics are used as error metrics in order to assess the difference between the two reanalysis datasets: Bias, Root-Mean-Square (RMSE), Scatter Index (SI), and Pearson's correlation coefficient (CorrCoeff).

Following the parametrization of the previous sections, the monthly values of these metrics are (1 stands for ERA5, and 2 for WW3):

$$\mathrm{Bias}(y,m,\cdot,\cdot) = \overline{(X_1 - X_2)}, \tag{18}$$

where $\overline{X} \equiv \mu_3(y,m,\cdot,\cdot)$ is calculated using Eq. 3.

$$\mathrm{RMSE}(y,m,\cdot,\cdot) = \sqrt{\overline{(X_1 - X_2)^2}}\,, \tag{19}$$

$$\mathrm{SI}(y,m,\cdot,\cdot) = \mathrm{RMSE}\;\Big/\;\overline{X_1}\,, \tag{20}$$

$$\mathrm{CorrCoeff}(y,m,\cdot,\cdot) = \frac{\overline{X_1\,X_2} - \overline{X_1}\;\overline{X_2}}{\sqrt{\left(\overline{X_1^2} - \overline{X_1}^2\right)\left(\overline{X_2^2} - \overline{X_2}^2\right)}}. \tag{21}$$

In addition, according to Stopa and Cheung (2014), the following error metrics have also been computed

$$\mathrm{NBias}(y,m,\cdot,\cdot) = \overline{(X_1 - X_2)}\;\Big/\;\sqrt{\overline{X_1^2}}\,, \tag{22}$$

$$\mathrm{NSTD}(y,m,\cdot,\cdot) = \mathrm{std}(X_2)\;/\;\mathrm{std}(X_1)\,, \tag{23}$$

$$\mathrm{CRMSE}(y,m,\cdot,\cdot) = \mathrm{std}(X_1 - X_2)\;/\;\mathrm{std}(X_1)\,, \tag{24}$$

where $\mathrm{std}(X) \equiv \sigma_3(y,m,\cdot,\cdot)$ is calculated using Eq. 4.

The mean monthly and/or annual variability of these indicators can be studied by applying appropriate averaging as in Eqs. 5~6 and/or Eqs. 10~11.

NUMERICAL RESULTS

Seasonal Analysis

Following the analysis procedure presented in the previous sections, the mean monthly variability is assessed on the basis of the mean monthly values $\widetilde{\mu}_3(m,\cdot,\cdot)$ and $\widetilde{\sigma}_3(m,\cdot,\cdot)$, calculated by Eqs. 5~6. According to these, there is a zonal distribution of the values with a distinct different behaviour between the northern and the southern hemisphere. Especially, around the equator there is a zone with the least variability in all months. It seems that the variability of mean values of ERA5 is a bit lesser than the corresponding of WW3 in both hemispheres.

In addition, the seasonal variability on a regional basis is studied by analysing separately the 13 subregions shown in Fig. 1. In Figs. 3, the mean monthly values of $\widetilde{\mu}_3(m,\cdot,\cdot)$ and $\widetilde{\sigma}_3(m,\cdot,\cdot)$ are depicted for 4 of the 13 subregions: two showing the variability in the northern hemisphere, and two in the southern. For comparison purposes, and apart from the two climatologies, the mean monthly values of satellite data are also plotted. As expected, the general picture shows more pronounced variability in the extratropical areas (ETNA, North Atlantic; ETNP, North Pacific; ETSP, South Pacific; ETSA, South Atlantic), and lesser one in the tropical zone (TWNP, TNAO, TESP, TSAO). Overall, there is better agreement between ERA5 and WW3, rather than with the satellite data; especially, in the extratropical subregions (ETNP, ETNA, ETSP, ETSA). In the tropical zone, satellite data are also in agreement with ERA5 and WW3 (TWNP, TNAO, TESP, TSAO). Generally, small deviations between satellite and model data can be attributed to the fact that the former have been estimated using 18 years (1992-2009), while the latter 31 years (1979-2009).

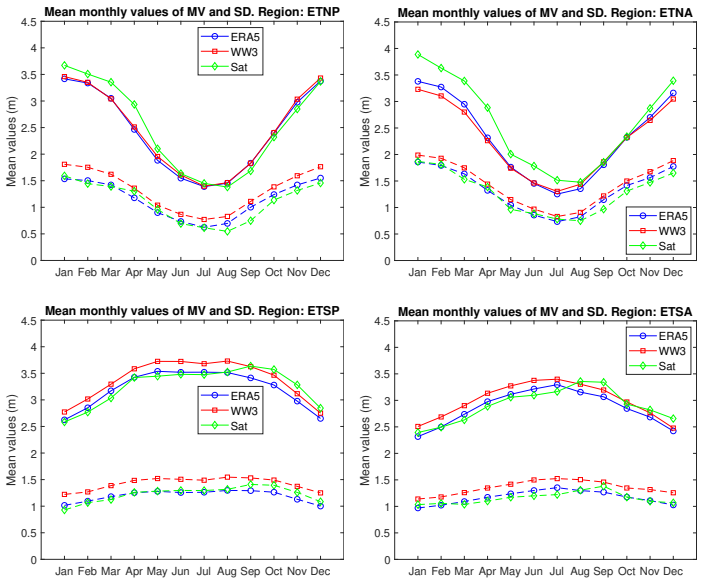

Fig. 3. Mean monthly values for northern (top: ETNP, ETNA) and southern hemisphere (bottom: ETSP,ETSA); see also Fig. 1 (continuous: mean value, dashed: standard deviation)

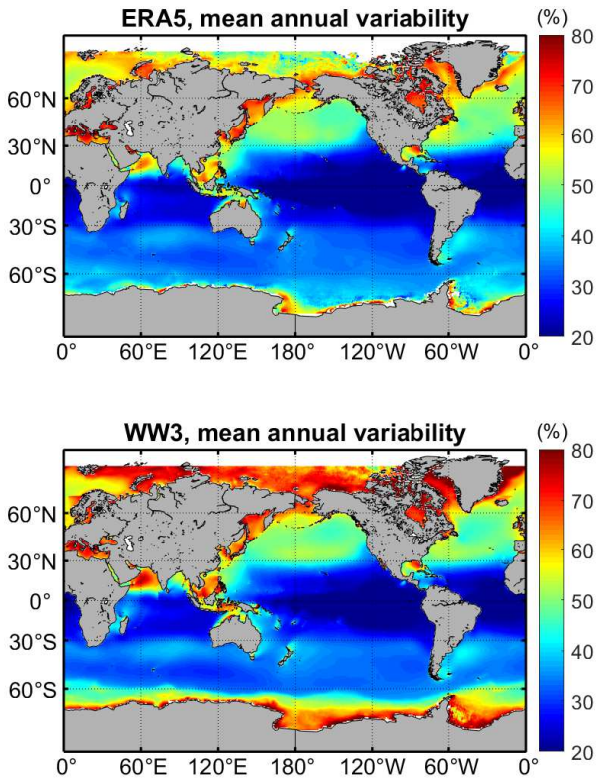

Fig. 4. Mean annual variability (in %) (top: ERA5, bottom: WW3)

Further, the mean annual variability is investigated in terms of $\widetilde{cv}_{32}(\cdot,\cdot)$; see Fig. 4. The agreement between ERA5 and WW3 datasets is good, showing more pronounced variability in the North Pacific and North Atlantic Ocean. Finally, the inter-annual variability is studied on the basis of $\widehat{cv}_{32}(\cdot,\cdot)$; see Eq. 13. In Fig. 5, the coefficient of variation is plotted for both datasets with WW3 exhibiting greater variability in most of the areas. It is worth noted that, the present WW3 results are in accordance with findings in Stopa et al. (2013).

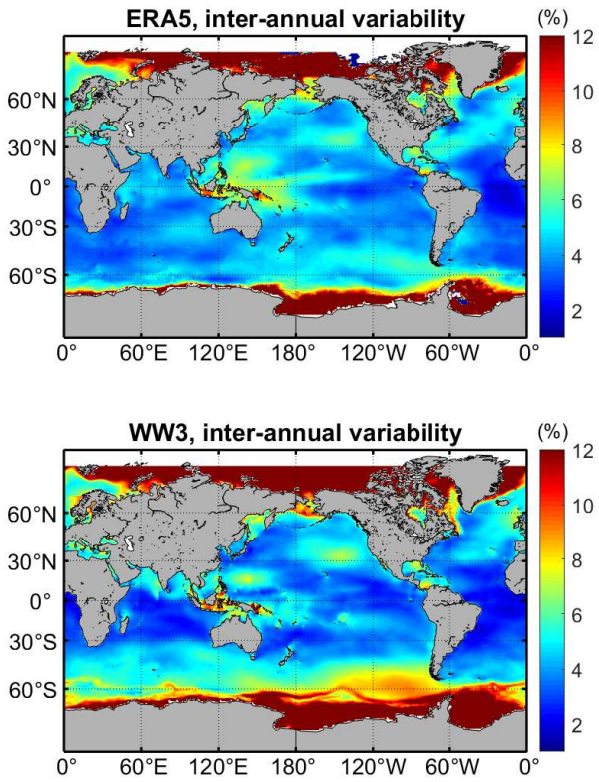

Fig. 5. Inter-annual variability (in %) (top: ERA5, bottom: WW3)

Probability Analysis

In this part of the analysis, the monthly empirical cdf's of ERA5 and WW3 are calculated, and, then, the quantiles of 50% (median), 90%, 95%, 99%, and 99.9% have been chosen to described the behaviour of the empirical distribution. Due to the fact that the distribution is not symmetric, the median describes better than the mean value the mean behaviour of the distribution. The other four are used for the description of the tail.

In Figs. 6–7, the mean annual fields of 50% and 99% quantiles is depicted for ERA5 and WW3. In general, there is again a zonal distribution of the quantiles, and ERA5 seems to exhibit less variability than WW3. The WW3 results are again in accordance with Stopa et al. (2013).

Further, in Figs. 8 the overall (averaged over all points) mean monthly values of the quantiles are plotted, giving a bird eye's view of the monthly variability of the distribution. It seems that WW3 distribution has higher values in all quantiles. In addition, ERA5 exhibits greater (month-to-month) variability in the 99.9% quantile.

In any case, the difference between the monthly values of the various quantiles are on the average in the range of ±15% of the WW3 values.

Error Analysis

In Figs. 9, the overall (averaged over all points) mean monthly variability of the error measures is given, namely Bias, RMSE, SI, CorrCoeff, and NBias, NSTD, CRMSE. Similarly, in Figs. 10, the overall (averaged over all points) mean annual variability of the same quantities is also plotted (bottom). Most of them seem to have no significant variability at all (except for the mean annual Bias). It is worth noticing that, CorrCoeff is near 1, which means that the two datasets are fully correlated.

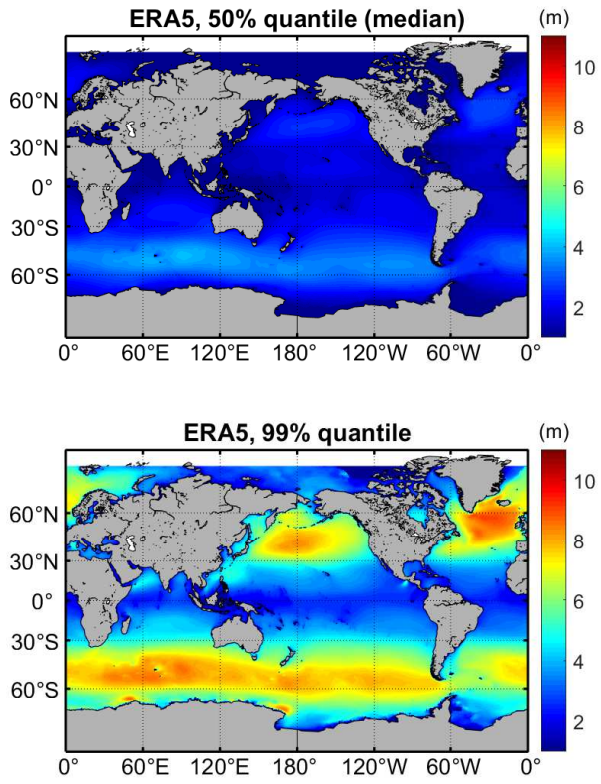

Fig. 6. Mean annual fields of ERA5 quantiles 50% and 99%

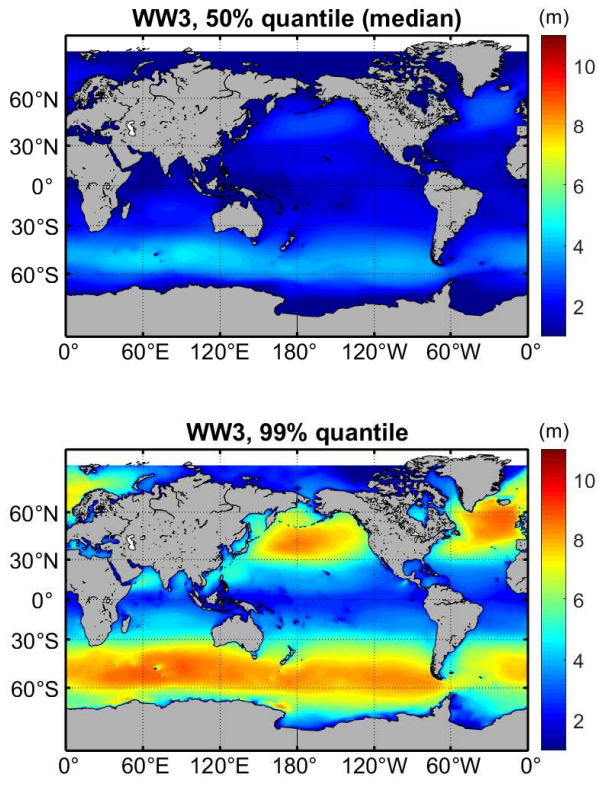

Fig. 7. Mean annual fields of WW3 quantiles 50% and 99%

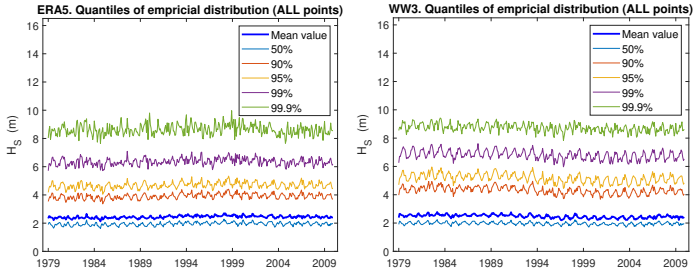

Fig. 8. Monthly quantiles 50% (median), 90%, 95%, 99%, and 99.9% (in m) (top: ERA5, bottom: WW3)

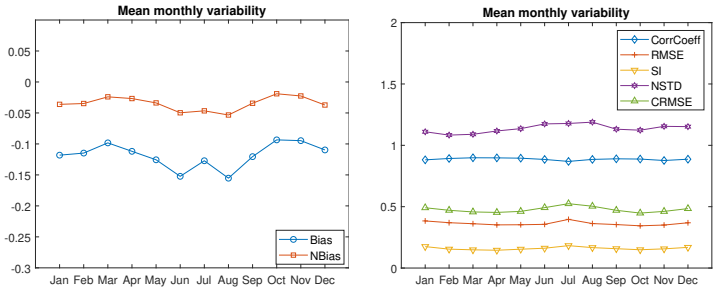

Fig. 9. Mean monthly variability of error measures

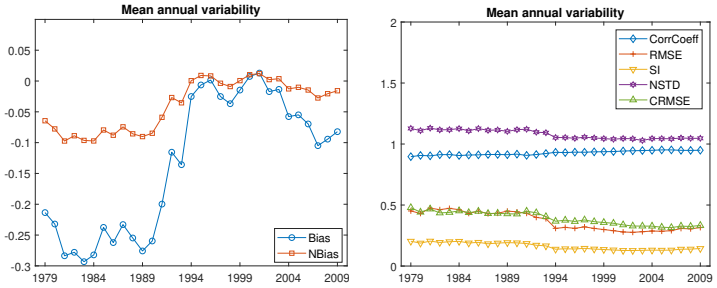

Fig. 10. Mean annual variability of error measures

Further, the overall (averaged over all the years) mean annual variability of the same measures is depicted in Figs. 11–12. One may argue that there is an inconsistency between the spatial distribution of the values of RMSE and those of CorrCoeff. However, the two error metrics should seen as complementary, giving only partially overlapping (and not exactly the same) statistical information. According to Taylor (2001), a more convenient measure to be related with CorrCoeff is the CRMSE, which is normalized with the standard deviation of the data; see Eq. 24. Indeed, one can observe that in the areas where CorrCoeff suggests low correlation between the two datasets, CRMSE error exhibits its highest values.

Finally, the same error measures have been calculated for the 13 subregions, using the monthly values of the datasets. The results both for a) ERA5 vs Satellite, b) WW3 vs Satellite are given in Table 1. In most of the regions, Bias is negative (satellite data have greater values than ERA5 and WW3), except for regions: TESP (ERA5 greater than satellite), and TSAO (both ERA5 and WW3 greater than satellite). The values of RMSE are pretty close for each area. As already mentioned, CorrCoeff is near 1, which is an indication of the correlation of the datasets. NSTD, which shows the ratio of the standard deviations of the two datasets, varies between 1.14–1.70 for "WW3 vs satellite" (satellite has greater variability 14–70%), and 1.16–1.97 for "ERA5 vs satellite" (satellite shows even greater variability compared to ERA5).

Table 1. Error measures between a) (E)RA5 and (S)atellite, b) (W)W3 and (S)atellite for the 13 subregions

	Bias		RMSE		SI		CorrCoeff		NBias		NSTD		CRMSE	
	E-S	W-S	E-S	W-S	E-S	W-S	E-S	W-S	E-S	W-S	E-S	W-S	E-S	W-S
ETNP	-0.18	-0.19	0.44	0.45	0.18	0.18	0.90	0.90	-0.07	-0.07	1.16	1.15	0.52	0.51
ETNA	-0.36	-0.42	0.59	0.65	0.25	0.28	0.88	0.88	-0.15	-0.18	1.28	1.38	0.62	0.70
TNIO	-0.08	-0.12	0.38	0.37	0.25	0.25	0.83	0.84	-0.05	-0.07	1.41	1.26	0.79	0.68
TWNP	-0.12	-0.16	0.33	0.35	0.19	0.21	0.74	0.75	-0.07	-0.09	1.34	1.32	0.89	0.87
TENP	-0.02	-0.07	0.21	0.23	0.10	0.11	0.78	0.77	-0.01	-0.03	1.28	1.23	0.80	0.78
TNAO	-0.09	-0.09	0.23	0.24	0.13	0.13	0.82	0.81	-0.05	-0.05	1.28	1.33	0.73	0.79
TSIO	-0.07	-0.11	0.28	0.30	0.13	0.14	0.79	0.78	-0.03	-0.05	1.36	1.26	0.83	0.78
TWSP	-0.13	-0.24	0.26	0.34	0.15	0.20	0.41	0.36	-0.08	-0.15	1.97	1.61	1.81	1.56
TESP	0.03	-0.03	0.32	0.31	0.14	0.14	0.52	0.55	0.01	-0.01	1.85	1.70	1.58	1.42
TSAO	0.01	0.02	0.26	0.26	0.14	0.14	0.61	0.59	0.01	0.01	1.76	1.59	1.39	1.28
ETSI	-0.17	-0.07	0.39	0.40	0.11	0.11	0.82	0.76	-0.05	-0.02	1.29	1.17	0.74	0.77
ETSP	-0.15	-0.12	0.35	0.39	0.11	0.12	0.76	0.69	-0.05	-0.03	1.31	1.14	0.85	0.85
ETSA	-0.15	-0.10	0.39	0.39	0.13	0.13	0.66	0.62	-0.05	-0.03	1.36	1.26	1.02	1.01

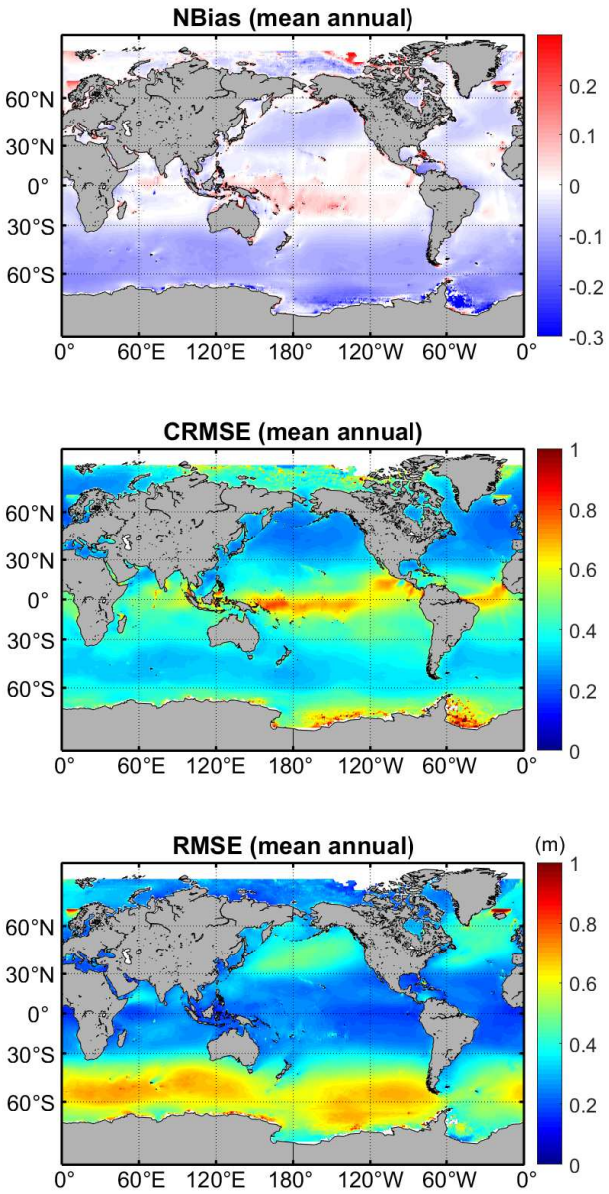

Fig. 11. Mean annual variability of NBias, CRMSE, RMSE (from top to bottom)

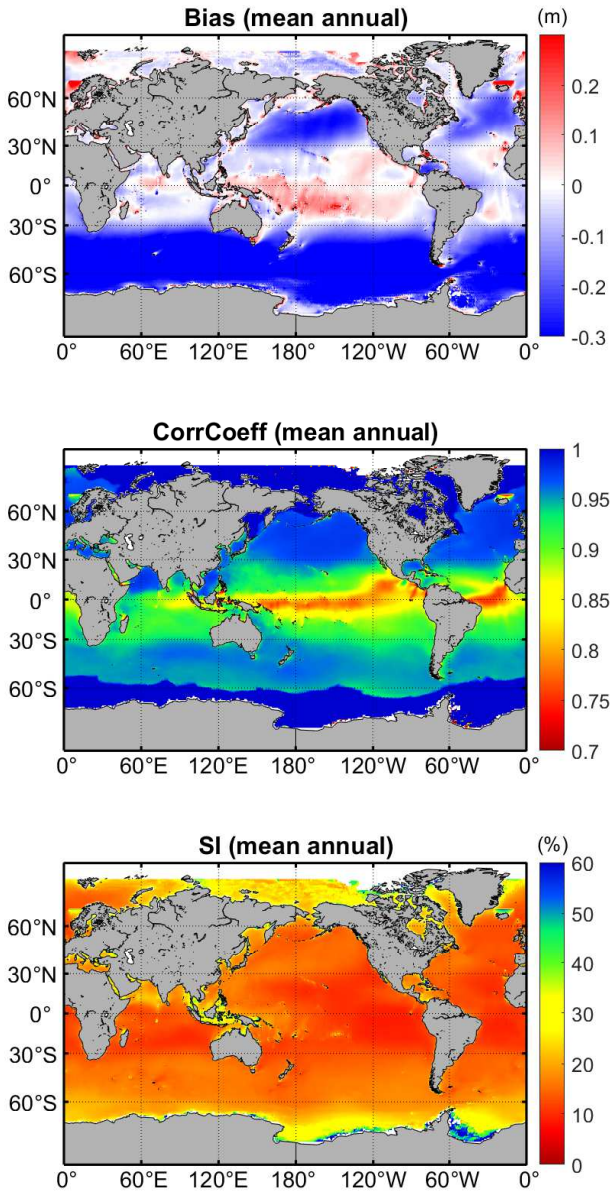

Fig. 12. Mean annual variability of Bias, CorrCoeff, SI (from top to bottom)

CONCLUDING REMARKS

In the present paper, the newly released ERA5 wave climatology is compared against WW3 climatology, as well as a merged satellite database. The intercomparison covers a period of 31 years (1979-2009), and has been performed using three-hourly wave fields for the entire globe.

The mean monthly values of the datasets are used as cornerstones of the analysis. First, seasonal analysis is performed based on seasonal characteristics of them showing a very good agreement. Moreover, the mean annual and the inter-annual variability computed are in accordance with findings of other researchers (Stopa et al., 2013).

Then, quantiles of the empirical cumulative distribution are calculated in order to get a picture of the variability of the distribution and of several percentiles of interest; especially in the tail of the distribution. Finally, several error measures are derived in order to assess the agreement between the datasets.

In addition to the field analysis, data are considered for thirteen non-overlapping subregions following Alves (2006), and the above mentioned analysis (seasonal, probability, error) is performed for each one of them.

All in all, the two datasets are in a very good agreement, with WW3 having little greater variability than ERA5. It is also noted that, wind speed statistics are also available, showing similar characteristics. However, they have been omitted from the present work due to space limitation.

ACKNOWLEDGMENTS

The present work has been performed in the framework of the research project "HDwave: High-dimensional statistical modelling of changes in wave climate and implications for maritime infrastructure" funded by the Norwegian Research Council under the contract No. 243814/E10.

REFERENCES

Alves JHG (2006). "Numerical modeling of ocean swell contributions to the global wind-wave climate," *Ocean Modelling*, 11(1), 98–122.

Athanassoulis G and Stefanakos Ch (1995). "A nonstationary stochastic model for long-term time series of significant wave height," *Journal of Geophysical Research Section Oceans*, 100(C8), 16149–16162.

Caires S, Sterl A, Bidlot JR, Graham N, and Swail V (2004). "Intercomparison of different wind wave reanalyses," *Journal of Climate*, 17(10), 1893–1913.

Campos RM and Guedes Soares C (2016). "Comparison and assessment of three wave hindcasts in the North Atlantic Ocean," *Journal of Operational Oceanography*, 9(1), 26–44.

Chawla A, Spindler DM, and Tolman HL (2013). "Validation of a thirty year wave hindcast using the Climate Forecast System Reanalysis winds," *Ocean Modelling*, 70, 189–206.

Copernicus Climate Change Service (C3S) (2017). "ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate," Copernicus Climate Change Service Climate Data Store (CDS), http://cds.climate.copernicus.eu, [accessed 2018.10–2019.03].

Cox AT and Swail VR (2001). "A global wave hindcast over the period 1958-1997: Validation and climate assessment," *Journal of Geophysical Research Oceans*, 106(C2), 2313–2329.

Guedes Soares C, Weisse R, Carretero JC, and Alvarez E (2002). "A 40 year hindcast of wind, sea level and waves in European Waters," in *21st International Conference on Offshore Mechanics and Arctic Engineering, OMAE'2002*, Oslo, Norway, pp. 669–675.

Gulev SK and Grigorieva V (2006). "Variability of the winter wind waves and swell in the North Atlantic and North Pacific as revealed by the Voluntary Observing Ship data," *Journal of Climate*, 19(21), 5667–5685.

Lefèvre JM and Cotton PD (2001). *Satellite Altimetry and Earth Sciences*, chapter 7. Ocean surface waves, Academic Press, pp. 305–328.

Perez J, Menendez M, and Losada IJ (2017). "GOW2: A global wave hindcast for coastal applications," *Coastal Engineering*, 124, 1–11.

Queffeulou P and Croizé-Fillon D (2017). *Global altimeter SWH data set*, Technical report, Laboratoire d' Océanographie Physique et Spatiale, IFREMER, Plouzané, France.

Rascle N and Ardhuin F (2013). "A global wave parameter database for geophysical applications. Part 2: Model validation with improved source term parameterization," *Ocean Modelling*, 70, 174–188.

Reguero BG, Menéndez M, Méndez FJ, Mínguez R, and Losada IJ (2012). "A Global Ocean Wave (GOW) calibrated reanalysis from 1948 onwards," *Coastal Engineering*, 65, 38–55.

Semedo A, Sušelj K, Rutgersson A, and Sterl A (2011). "A global view on the wind sea and swell climate and variability from ERA-40," *Journal of Climate*, 24(5), 1461–1479.

Stefanakos Ch, Athanassoulis G, and Barstow S (2006). "Time series modeling of significant wave height in multiple scales, combining various sources of data," *Journal of Geophysical Research Section Oceans*, 111(C10), 10001–10012.

Stefanakos Ch and Schinas O (2014). "Forecasting bunker prices; a non-stationary, multivariate methodology," *Transportation Research Part C Emerging Technologies*, 38(1), 177 – 194.

Sterl A and Caires S (2005). "Climatology, variability and extrema of ocean waves: The web-based KNMI/ERA-40 wave atlas," *International Journal of Climatology*, 25(7), 963–977.

Stopa JE (2018). "Wind forcing calibration and wave hindcast comparison using multiple reanalysis and merged satellite wind datasets," *Ocean Modelling*, 127, 55 – 69.

Stopa JE and Cheung KF (2014). "Intercomparison of wind and wave data from the ECMWF Reanalysis Interim and the NCEP Climate Forecast System Reanalysis," *Ocean Modelling*, 75, 65 – 83.

Stopa JE, Cheung KF, Tolman HL, and Chawla A (2013). "Patterns and cycles in the Climate Forecast System Reanalysis wind and wave data," *Ocean Modelling*, 70, 207 – 220, ocean Surface Waves.

Taylor KE (2001). "Summarizing multiple aspects of model performance in a single diagram," *Journal of Geophysical Research Atmospheres*, 106(D7), 7183–7192.

Tolman HL, Balasubramaniyan B, Burroughs LD, Chalikov DV, Chao YY, Chen HS, and Gerald VM (2002). "Development and implementation of wind-generated ocean surface wave models at NCEP," *Weather and Forecasting*, 17(2), 311–333.

The WAMDI Group (1988). "The WAM model-A third generation ocean wave prediction model," *Journal of Physical Oceanography*, 18(12), 1775–1810.

Young IR, Zieger S, and Babanin AV (2011). "Global trends in wind speed and wave height," *Science*, 332(6028), 451–455.